Geocoding Administrative Files at ICBS

Anna Beenstock Cohen

GIS Sector

May 2016

What will we talk about

- Geocoding at ICBS
- ICBS geocoding and symbolizing concept, system & main processes
- Demonstrating the process with:

- Population Register
- Building & Dwelling Register

What is it all about?

- Administrative files
- Quality & usage variations.
- Field work context: sampling for surveys
- Examples: population register, property sales, business register, pupils file etc.

Defining Geocoding

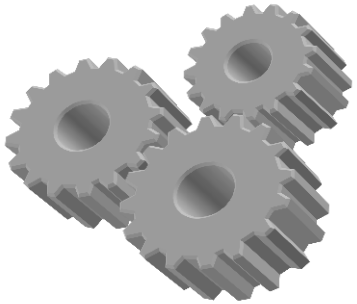
Associating an item to its exact spatial location based on its address attributes

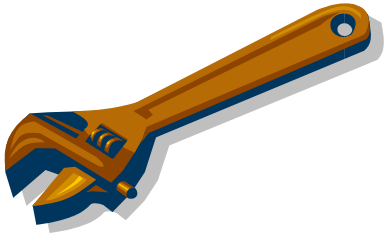
The goal:

To provide the most exact XY location possible for every record.

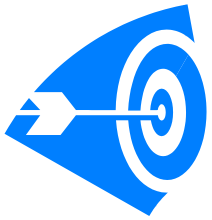
Geocoding provides the geographical context of the data

Geocoding: Automatic & Manual





Automatic	Manual
fast ↑	slow ↓
Less Accurate ↓	Accurate ↑



Geocoding and mapping challenges

The means

- Methods and Sources for Geo-Coding: Exact match, Interpolation, Combined.
- Geo-Coding Phases:
  - Records with full addresses

  - Records with a partial address
  - Records without Addresses

  - Levels of Geocoding

Proceeding Acts for Automatic Generic Application of Geo-Coding

1. Create Tables by Spatial Analysis of GIS layers (using SQL Geometry).
2. The Application uses Stored procedures.
3. The application preforms modules of the whole process from reading the file to giving each record its coordinates.
4. The application outputs are coordinates of a building centroid or a statistical area centroid.

Israeli Population Register

- Received: 3 times a year at ICBS, arrives coded
- Source: Population & Immigration Authority
- Geographical coverage: National
- Description: 9,116,537 records
- Sampling frame for surveys: Social Survey, Longitude Survey

Israeli Population Register by Geocode status

Geocode Data Source	Geocode Accuracy	Population	Population Percent	Addresses	Addresses Percent
Exact match - addresses	High	5,980,071	71.5	374,837	66.6
Interpolation of Bedouin land plots	High	109,810	1.3	8,141	1.4
Center of settlement with one statistical area	Medium	1,145,517	13.6	91,106	16.1
Statistical area from none Jewish neighborhoods	Medium	50,404	0.6	747	0.1
Other interpolations		1,076,103	12.8	87,980	15.6

- Total Addresses : 571,385
- Addresses geocoded: 563,811 (98.6%)

Building and Dwelling Register

- Received: Annually, May-June
- Source: 10 different suppliers that give service to 1,214 localities in Israel for property tax services
- Geographical coverage: per locality & supplier
- Description: 3,749,181records in 2015
- Sampling frame for surveys: Labour Force Survey and Household Expenditure Survey

Building and Dwelling Register (Cont.)

- Assets: Classified by dwelling (67%) and other purposes (commerce, education, etc.)
- Geocoding via 2 paths results for dwellings:

Path	Geocoded to building (percentage)	Geocoded to Statistical Area (percentage)
Address	79.1	92.5
Parcel	60.4	65

- Algorithm that choses for each statistical area by the higher quality output.

Building and Dwelling by Geocode status

Geocode Data Source	Geocode Accuracy	Addresses	Addresses Percent
Exact match - addresses	High	2,229,281	59.4
land Interpolation of Bedouin plots	High	13,415	0.3
Center of settlement with one statistical area	Medium	472,424	12.6
Statistical area from none Jewish neighborhoods	Medium	11,883	0.3
Other interpolations		433,217	11.5

- Total Addresses: 3,749,181
- Addresses geocoded: 3,160,220

Building and Dwelling by Geocode status

Geocode Data Source	Geocode Accuracy	Parcel	Parcel Percent
Exact match - One building in parcel	High	1,154,463	30.7
Chosen building from parcel	High	672,997	17.9
Statistical area from parcel	Medium	238,840	6.3

To Sum Things up

- The connection between administrative files, geocoding and sampling
- Geocoding challenges have been translated into procedures and algorithms at ICBS
- Demonstrating process with:

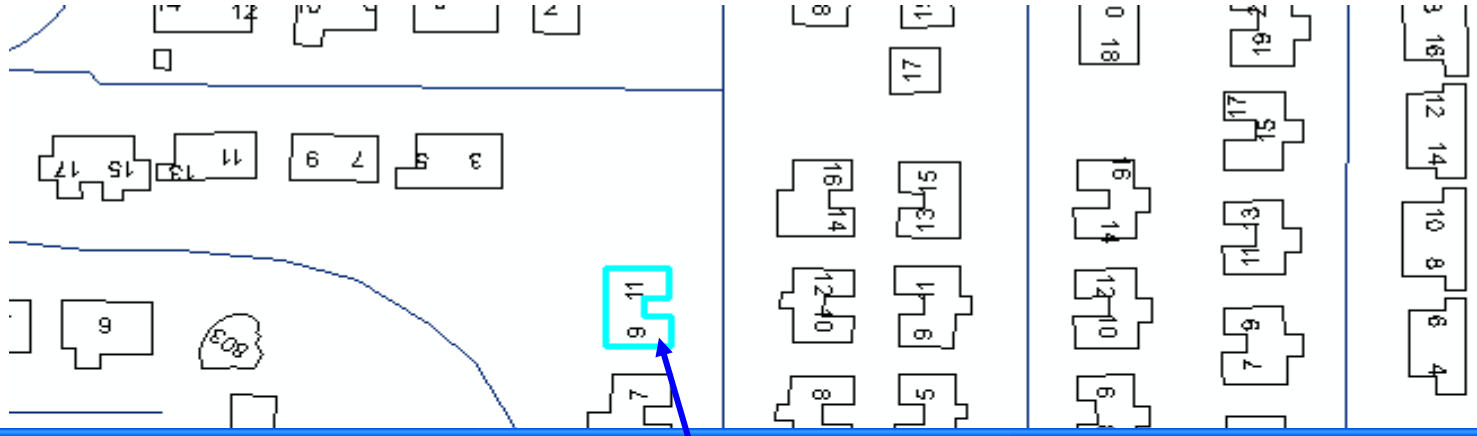
- Population Register
- Building & Dwelling Register

EU TWINNING PROJECT 2016-2018
ENHANCING THE QUALITY OF OFFICIAL STATISTICS

הלשכה המרכזית לסטטיסטיקה
Central Bureau of Statistics
دائرة الإحصاء المركزية

Thank You for your time and always remember:

GIS

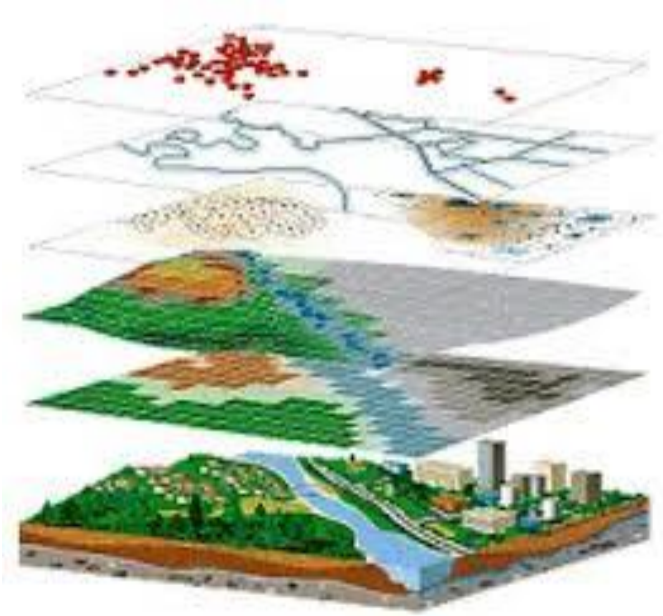


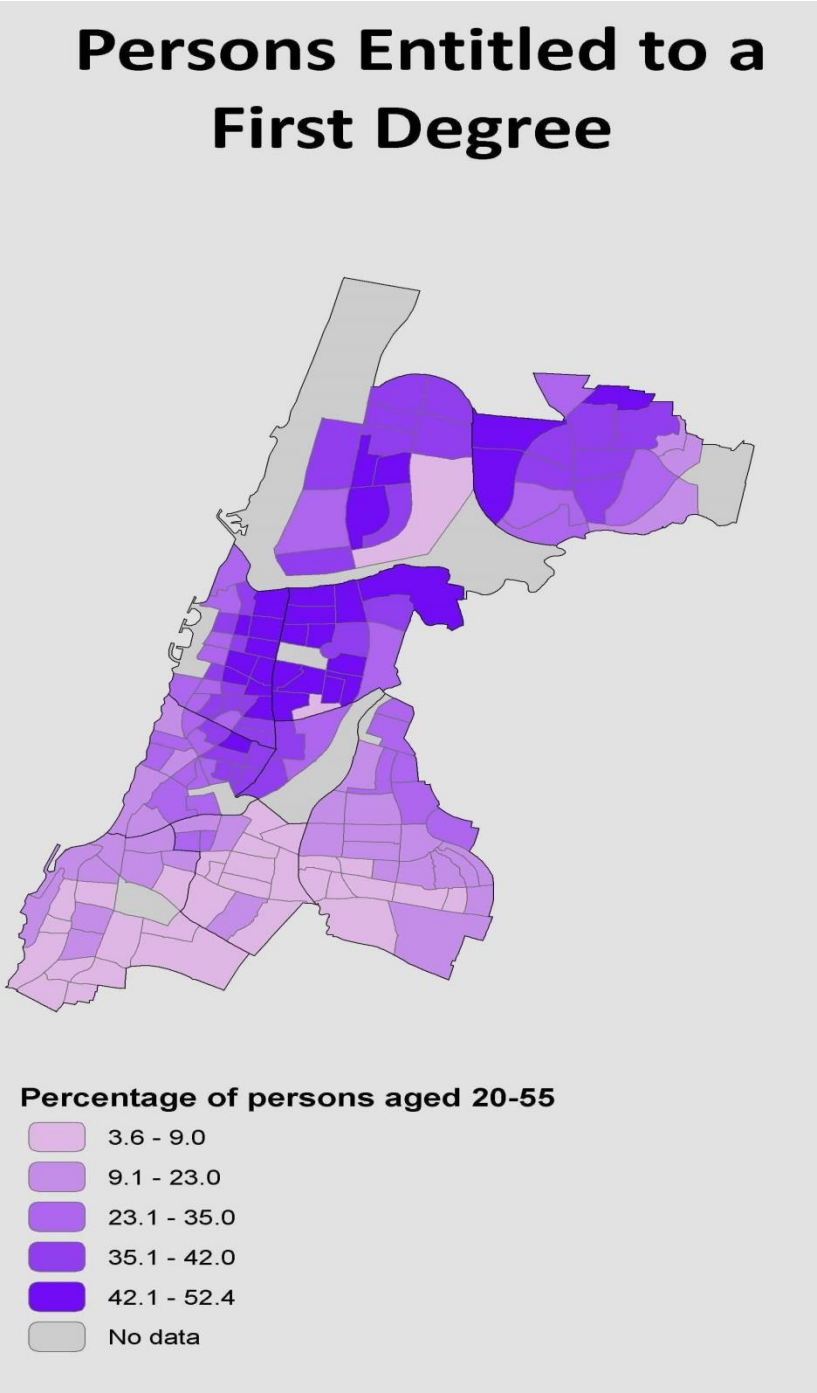
Attributes of EYALM.SDEBuildings

*OBJECTID	TYPE_CODE	HEIGHT	LND_HGT	SEMEL_YISHUV	FLAG_FIC	MEZAHE	UNIQUE7_ID	*Shape	SHAPE.AREA	SHAPE.LEN
14526	701	270.66	0	1200	<Null>	<Null>	1128871	Polygon	199.068406	69.041450
1055	701	271.7	0	1200	<Null>	<Null>	1128872	Polygon	208.062269	62.124248
17821	701	273.01	0	1200	<Null>	<Null>	1128873	Polygon	187.548896	64.449376
22388	701	274.3	0	1200	<Null>	<Null>	1128874	Polygon	154.543004	55.310039
23449	701	273.8	0	1200	<Null>	<Null>	1128875	Polygon	162.229865	51.046973
380	701	273.32	0	1200	<Null>	<Null>	1128876	Polygon	162.064568	53.081514
3513	701	277.11	0	1200	<Null>	<Null>	1128877	Polygon	305.087654	82.749787
8665	701	272.62	0	1200	<Null>	<Null>	1128878	Polygon	299.254837	74.894827
2003	701	278.34	0	1200	<Null>	<Null>	1128879	Polygon	293.836250	75.627362
3954	701	282.36	0	1200	<Null>	<Null>	1128880	Polygon	314.306931	84.753425
6493	701	289.45	0	1200	<Null>	<Null>	1128881	Polygon	296.234934	92.155423
7620	701	289.89	0	1200	<Null>	<Null>	1128882	Polygon	295.342092	85.735720
23669	701	290.51	0	1200	<Null>	<Null>	1128883	Polygon	258.447194	88.547051
29289	701	289.3	0	1200	<Null>	<Null>	1128884	Polygon	300.552025	96.241585
2192	701	276.08	0	1200	<Null>	<Null>	1128885	Polygon	80.908950	37.620226

GIS Data for Geocoding

- Addresses
- Buildings
- Streets
- Municipal borders
- Neighborhoods
- Statistical areas:





Definition of statistical areas

Urban localities with more than 10,000 residents are divided into statistical areas. These are small geographic units as homogeneous as possible, which it is possible to reflect the unique characteristics of certain areas within a locality. A residential statistical area generally numbers 3,000 – 4,000 persons.

Coding

Changing text to codes for localities and streets based on dictionaries

- Locality Name Formal Codes
- Street Name Formal Codes
- Codes for Synonym Names

synonyms

- ✓ Variations of Names
- ✓ Different statuses

Parsing

- ✓ String parsing to: Street Name, House Number, Entrance, Flat Number, Others

Parsing example: Pharmacies

בית מרקחת מכבי
טלפון: 08-8549503 |
כתובת: השייטים 12 אזור: אשדוד
שעות פעילות - א'-ה': 8:00-19:00
ו': 8:00-13:00

אסיא - ישועת חיים
טלפון: 08-8658828 | פקס: 08-8659793 |
כתובת: רבי אליעזר בן הורקנוס 22 אזור: אשדוד
מקבלים מרשמים של קופות: לאומית, מכבי, מאוחדת וכללית.

בית מרקחת תל חי
טלפון: 08-86768795 | 050-7108861 | פקס: 08-8678795 |
כתובת: שדרות תל חי 63/17 אזור: אשדוד
בית מרקחת שכונתי מכבד מרשמים של קופ"ח מכבי , מאוחדת און ליין ומרשמי כללית ממרפאת י"ד ובנוסף ניתן למצוא מוצרי טבע והומאופתיה. בבית המרקחת ניתן להכין הכנות רוקחיות כמעט מכל הסוגים.

בית מרקחת - דרך המלך
טלפון: 08-8651262, 052-7607080 | פקס: 08-8656660 |
כתובת: הנביאים אזור: אשדוד

מפרשי ארגמן
טלפון: 08-8670632, 050-5658579 | פקס: 013
כתובת: המלך דוד 18 אזור: אשדוד

שופרסל דיל אשדוד
טלפון: 08-8634400 |
כתובת: הבושם 7 אזור: אשדוד

דן קניון טי-מול
טלפון: 08-8554289, 08-8653877 |
כתובת: אחוזת בית 1, רובע ט' אזור: אשדוד

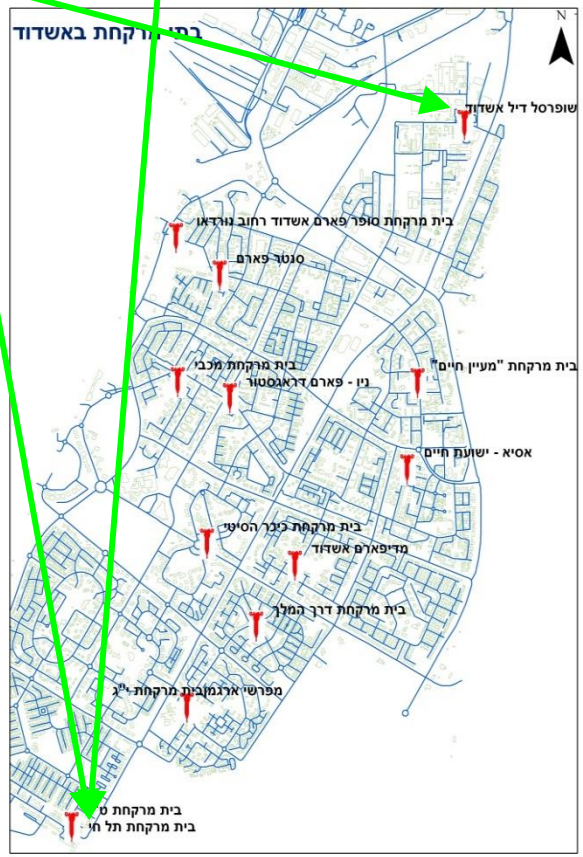
סנטר פארם
טלפון: 08-8524105 |
כתובת: רוגוזין 1 אזור: אשדוד

סופר פארם ז'בוטינסקי - אשדוד
טלפון: 08-8638000 |
כתובת: ז'בוטינסקי 41 אזור: אשדוד

סופר פארם הבנים אשדוד
טלפון: 08-8652062 |
כתובת: הבנים 1 אזור: אשדוד

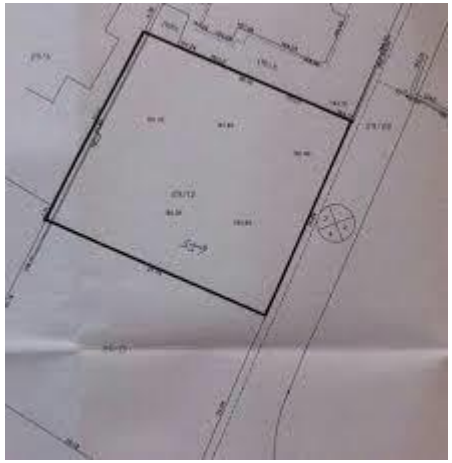
NYB	HOUSE_NO	STREET	NAME	CITY	ID
	12	השייטים	בית מרקחת מכבי	אשדוד	1
	22	רבי אליעזר בן הורקנוס	אסיא - ישועת חיים	אשדוד	2
	63	שדרות תל חי	בית מרקחת תל חי	אשדוד	3
		הנביאים	בית מרקחת - דרך המלך	אשדוד	4
	18	המלך דוד	מפרשי ארגמן	אשדוד	5
	7	הבושם	שופרסל דיל אשדוד	אשדוד	6
		רובע ט'	דן קניון טי-מול	אשדוד	7
	1	רוגוזין	סנטר פארם	אשדוד	8
	41	ז'בוטינסקי	סופר פארם ז'בוטינסקי - אשדוד	אשדוד	9
	1	הבנים	סופר פארם הבנים אשדוד	אשדוד	10
	1	יהודה הלוי	ניו - פארם דראגסטור	אשדוד	11
	1	הרצל פינת דרך מנחם בג	ניו - פארם	אשדוד	12
	20	שדרות הפרחים	מדיפארם אשדוד	אשדוד	13
	14	שבט לוי	מג'יק פארם	אשדוד	14
שכונה יא	522	כנרת	היפר פארם אמפורה	אשדוד	15
	7	מורדאו	בית מרקחת סופר פארם אשדוד	אשדוד	16
	20	הפלמ"ח	בית מרקחת "מעיין חיים"	אשדוד	17
	87	העצמאות	בית מרקחת כיכר הסיטי	אשדוד	18
	18	המלך דוד	בית מרקחת י"ג	אשדוד	19
	63	שדרות תל-חי	בית מרקחת ט"ו	אשדוד	20
	2	הנביאים	בית מרקחת דרך המלך	אשדוד	21

SmlYishuv	SmlKalfi	SmlRechov	ShemRechov	MspBayit	total	STREET_ID	OID	UNIQUE7_ID	CITY_STAT08
70		678	שד תל חי	63		700678	20	5082800	700433
70		678	שד תל חי	63		700678	3	5082800	700433
70		801	המלך דוד	18		700801	19	5082056	700411
70		801	המלך דוד	18		700801	5	5082056	700411
70		704	הנביאים	2		700704	21	5081415	700343
70		307	שדרות הפרחים	20		700307	13	111707	700321
70		852	העצמאות	87		700852	18	1207842	700317
70		432	השייטים	12		700432	1		700243
70		217	יהודה הלוי	1		700217	11	5079177	700234
70		753	אליעזר בן הורקנוס	22		700753	2	5079877	700212
70		207	הפלמ	20		700207	17	5079036	700143
70		128	רוגוזין	1		700128	8	5063375	700126
70		176	[illegible]	[illegible]		[illegible]	16		700122
70		507	[illegible]	[illegible]		[illegible]	6		700112
							0	0	0

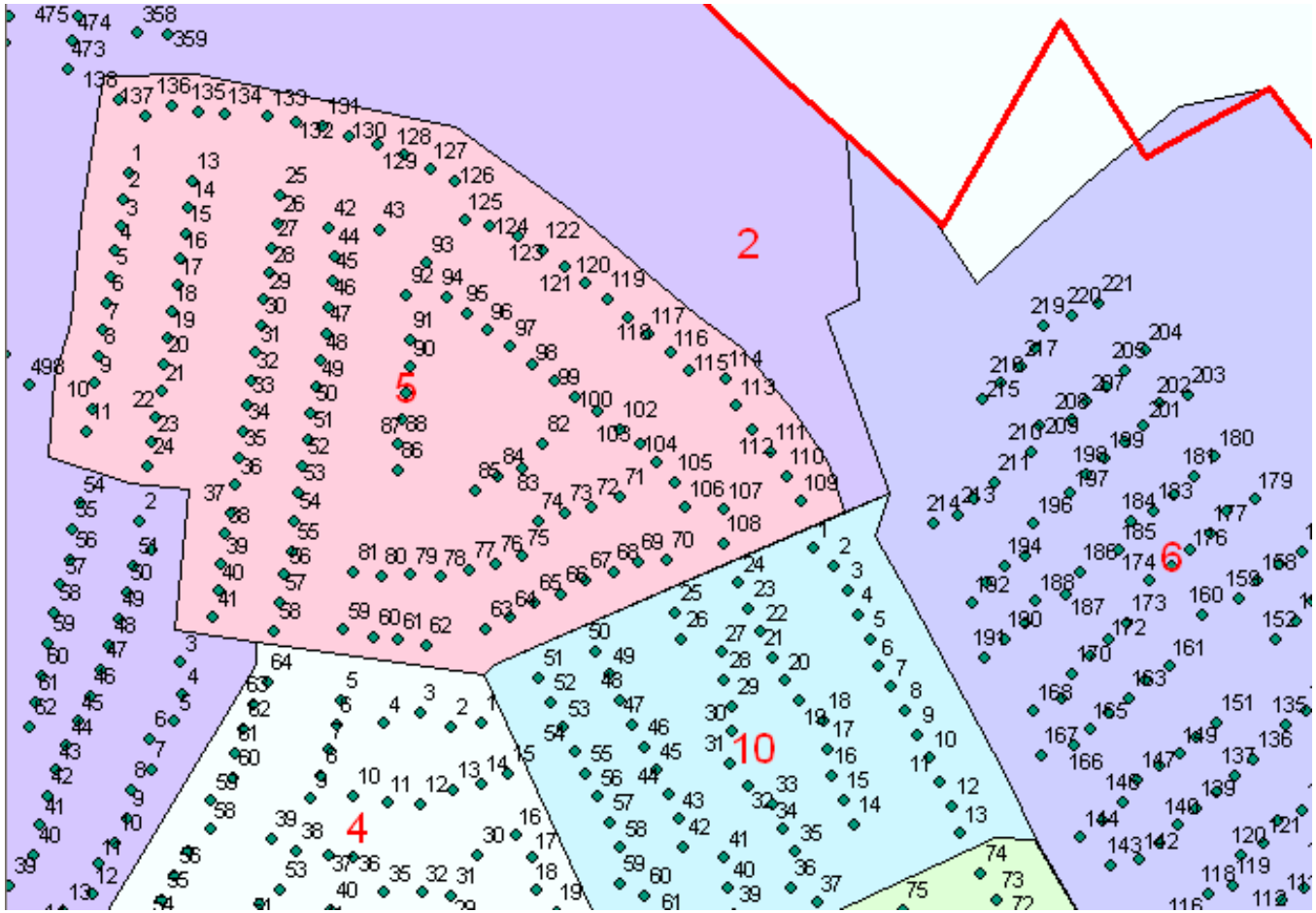


Geocoding by other spatial attributes

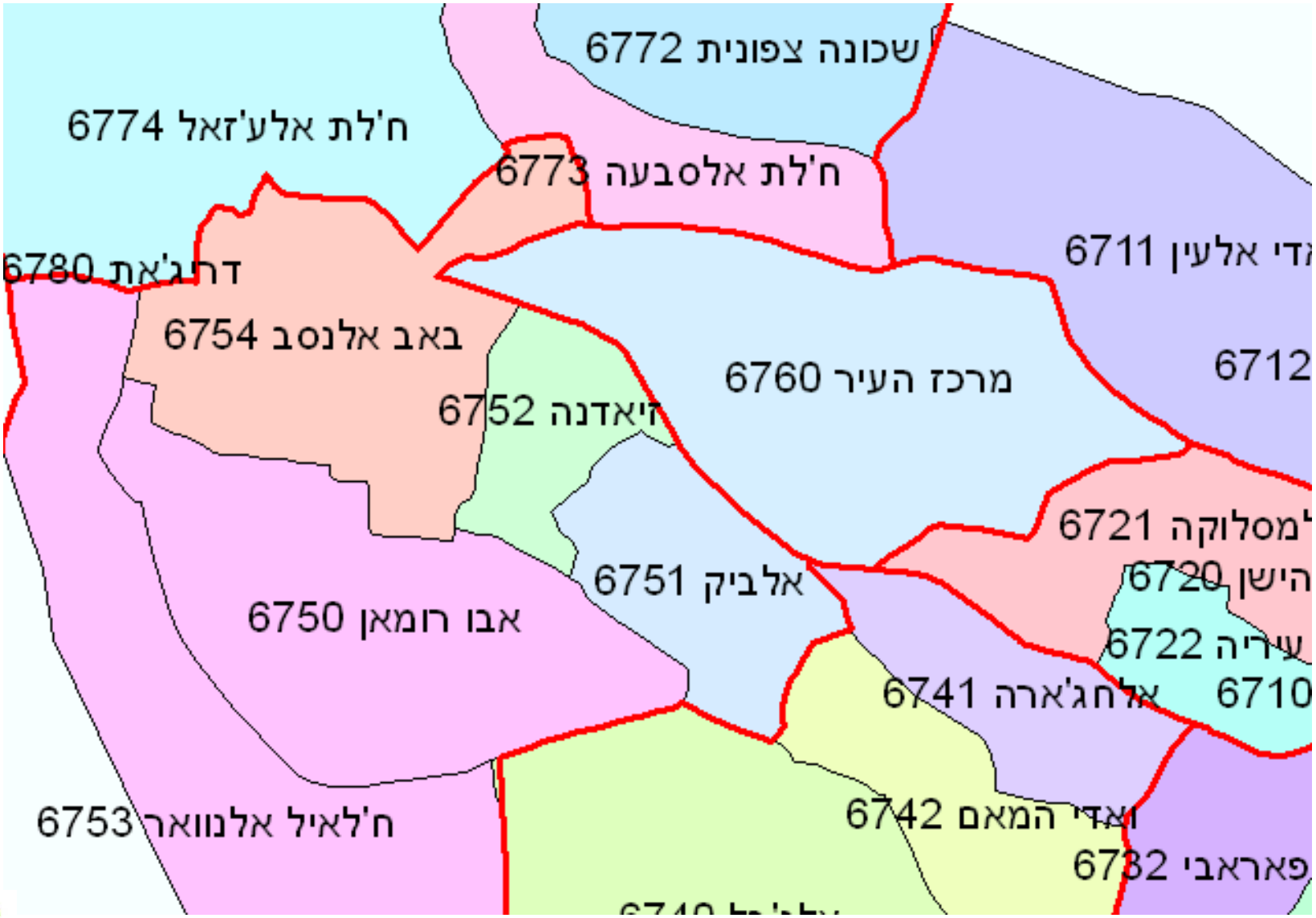
- ✓ Bedouin Localities
- ✓ Non-Addressed Localities
- ✓ Parcels



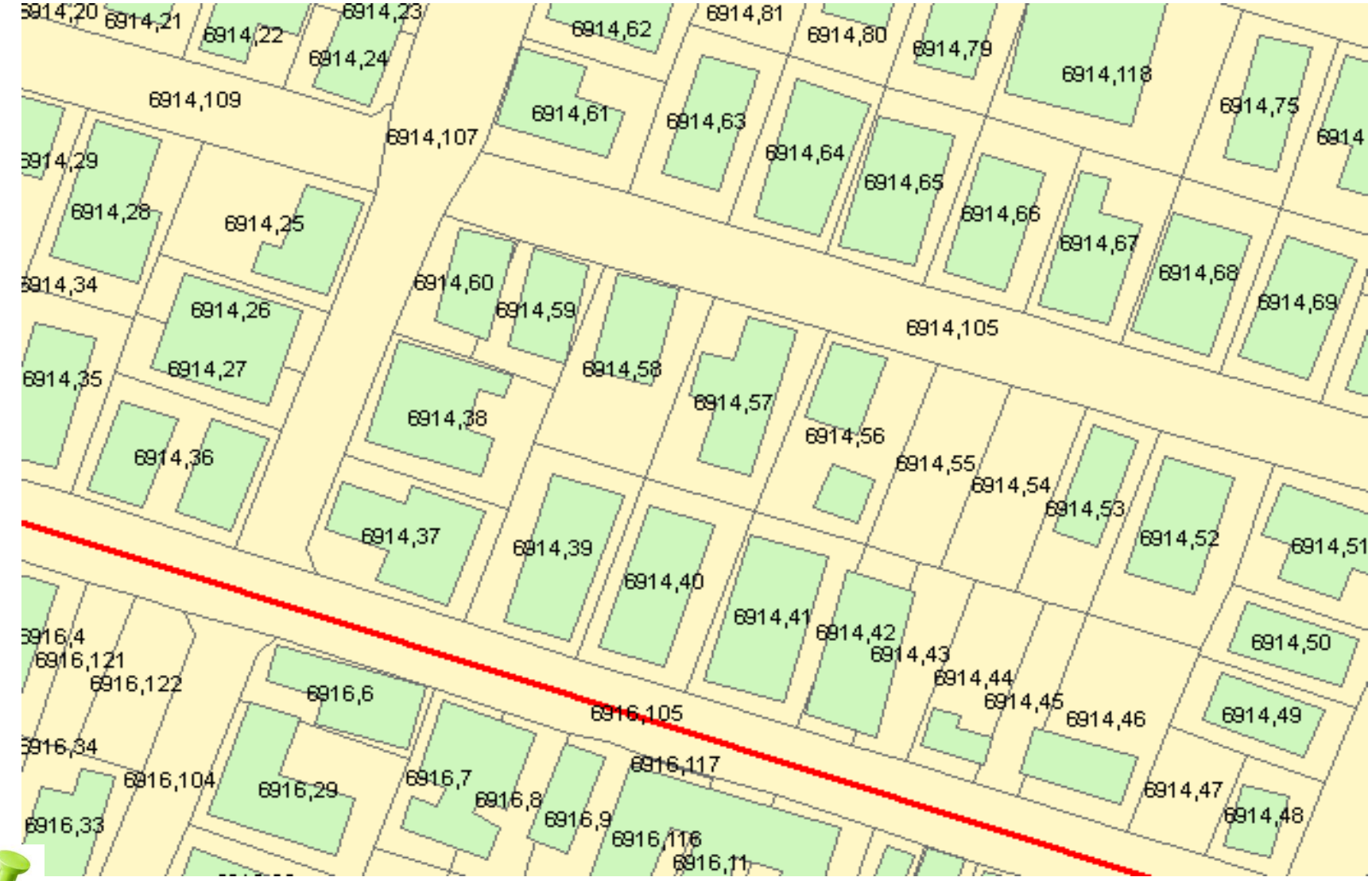
Land Plots in Bedouin Municipalities



Neighborhoods in Non Jewish Localities



Parcels from National Cadaster



Geocoding status

The system gives each record a status that indicates the geocode level of spatial accuracy reached.

Examples:

Data source	Accuracy	Status	Code
Addresses	high	Exact match	2001
One building in parcel	high	Exact match	2002
Chosen building from parcel	medium	Refined match	2003
Building 4 building away in addresses	medium	neighbors	2007
Statistical area from postal id's	low	Interpolation	3008
Statistical area from landmark table in none-Jewish settlement	low	Interpolation	3009